Combining diversity within Sorghum bicolor for genomic and fine mapping of intra-allelic interactions underlying heterosis

Fridman, E. The Hebrew University of Jerusalem
Yu, J. Iowa State University IA
Elbaum, R. The Hebrew University of Jerusalem

Project award year: 2012
Three year research project

1. Abstract

Heterosis, the enigmatic phenomenon in which whole genome heterozygous hybrids demonstrate superior fitness compared to their homozygous parents, is the main cornerstone of modern crop plant breeding. One explanation for this non-additive inheritance of hybrids is interaction of alleles within the same locus. This proposal aims at screening, identifying and investigating heterosis trait loci (HTL) for different yield traits by implementing a novel integrated mapping approach in *Sorghum bicolor* as a model for other crop plants. Originally, the general goal of this research was to perform a genetic dissection of heterosis in a diallel built from a set of *Sorghum bicolor* inbred lines. This was conducted by implementing a novel computational algorithm which aims at associating between specific heterozygosity found among hybrids with heterotic variation for different agronomic traits. The initial goals of the research are: (*i*) Perform genotype by sequencing (GBS) of the founder lines (*ii*) To evaluate the heterotic variation found in the diallel by performing field trails and measurements in the field (*iii*) To perform QTL analysis for identifying heterotic trait loci (HTL) *(iv)* to validate candidate HTL by testing the quantitative mode of inheritance in F2 populations, and *(v)* To identify candidate HTL in NAM founder lines and fine map these loci by test-cross selected RIL derived from these founders.

The genetic mapping was initially achieved with app. 100 SSR markers, and later the founder lines were genotyped by sequencing. In addition to the original proposed research we have added two additional populations that were utilized to further develop the HTL mapping approach; (1) A diallel of budding yeast (*Saccharomyces cerevisiae*) that was tested for heterosis of doubling time, and (2) a recombinant inbred line population of *Sorghum bicolor* that allowed testing in the field and in more depth the contribution of heterosis to plant height, as well as to achieve novel simulation for predicting dominant and additive effects in tightly linked loci on pseudo-overdominance.

There are several conclusions relevant to crop plants in general and to sorghum breeding and biology in particular: (i) heterosis for reproductive (1), vegetative (2) and metabolic phenotypes is predominantly achieved via dominance complementation. (ii) most loci that seems to be inherited as overdominant are in fact achieving superior phenotype of the heterozygous due to linkage in repulsion, namely by pseudo-overdominant mechanism. Our computer simulations show that such repulsion linkage could influence QTL detection and estimation of effect in segregating populations. (iii) A new height QTL (qHT7.1) was identified near the genomic region harboring the

known auxin transporter Dw3 in sorghum, and its genetic dissection in RIL population demonstrated that it affects both the upper and lower parts of the plant, whereas Dw3 affects only the part below the flag leaf. (iv) HTL mapping for grain nitrogen content in sorghum grains has identified several candidate genes that regulate this trait, including several putative nitrate transporters and a transcription factor belonging to the no-apical meristem (NAC)-like large gene family. This activity was combined with another BARD-funded project in which several de-novo mutants in this gene were identified for functional analysis.

Summary Sheet

PubType	IS only	Joint	US only
Review Article	1	0	0
Reviewed	1	1	0
Submitted	1	0	0

Mode of collaboration

The collaboration between two labs included exchange of students and visit of the Israeli PI in ISU, both for giving a departmental seminar on Heterosis (topic of the proposal), as well as coordinating the research in both labs. In addition, several Skype meetings were held every year in the presence of the PhD and postdocs involved in the project. In these meetings issues related to the set-up of the field experiments and the genome scan for associating heterosis with particular allelic variation were discussed. The US partner was also involved in resequencing of DNA libraries used for identifying mutants in genes of interest that were identified in the diallel analysis in Israel. Both PIs were invited to present the work of the project in several international and national meetings including the Plant and Animal Genome (PAG) at San-Diego. Imri Ben-Israel, a phd student in the lab was awarded a travel fellowship for BARD, which allowed him to travel to the Yu lab and refine the computational algorithm used for the genome scans (the heterotic trait loci [HTL] mapping). Genotyping of the Israeli material by resequencing was conducted at the US lab, and the analysis was conducted in the Israeli lab.

2. Achievements

2.1 Development of the heterotic trait loci (HTL) mapping in a diallel

The HTL mapping evolved from single point (marker) analysis (Figure 1; conducted with the sorghum SSR and SNP date) to a more sophisticated haplotype-based genome scan conducted in the fully sequenced *Saccharomyces cerevisiae* dialle (Figure 1D-G)).

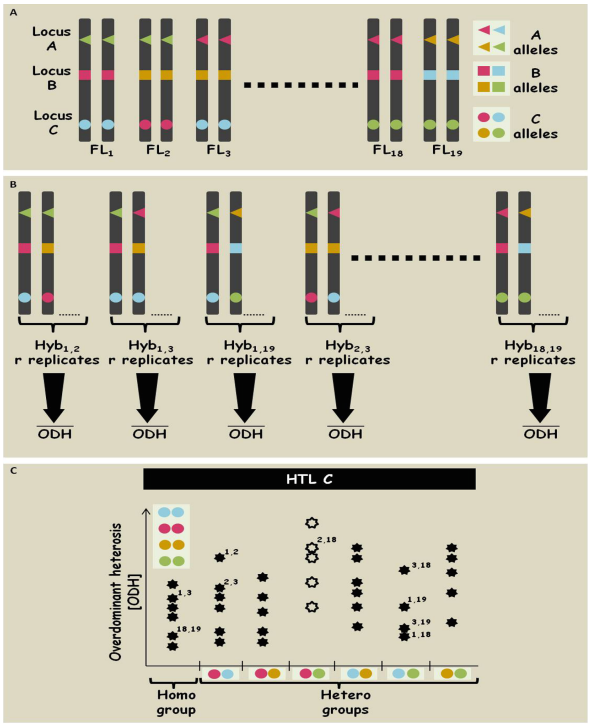

Figure 1. The heterotic trait loci (HTL) mapping with single point marker analysis and haplotype based mapping. **A.** Genotyping of the selected FLs is represented by three loci with different shapes (A, B and C), with each harboring 4 different alleles among the FLs of the diallel. **B.** Projection of the FLs genotype to the derived hybrids and calculation of the mean heterosis values (ODH) for the r replicates of an hybrid. **C.** Statistical analysis to identify specific hetero combinations with advantageous ODH values (the purple/green hetero-group in this illustration). Taken from Ben-Israel et al. 2012 (1).

For a fully sequenced exome, which has become a more relevant platform for plant breeding, we had to develop a methodology that will capture the haplotype structure in the founder lines of the diallel. We have chosen to follow a 2x3 rule that will allow minimal representation of 3 hybrids for each of the three possible genotypic groups, therefore allowing statistical comparison of their heterotic phenotype (HP or formerly overdominant heterosis, ODH) (Figure 2). According to the "2x3 rule" a variable sliding window is defined such that at least two alleles (haplotypes), each represented by at least three parents, are identified. This is with minimum 3 and maximum 10 SNP per haplotype. Such arrangement achieve 3 hybrids with identical local homozygousity for either one of the two alleles. We view this as a minimal requirement to compare means by a-parametric (e.g., Kruscal-Wallis) analysis as the HP values is not necessarily following a normal distribution. In addition to this change in the genotype to phenotype analysis an additional improvement was included that should increase the likelihood of identifying loci in which allelic interactions are affecting the non-additive inheritance in trios of two inbreds and one hybrid. This step includes comparison of the HP values between the two homozygous groups (Figure 2E); if these values are not significantly different then we assume

that effects on HP in this locus (identified in the first step of genome scan; Figure 2D). The manuscript describing this improvement is currently under review in the GENOME journal.

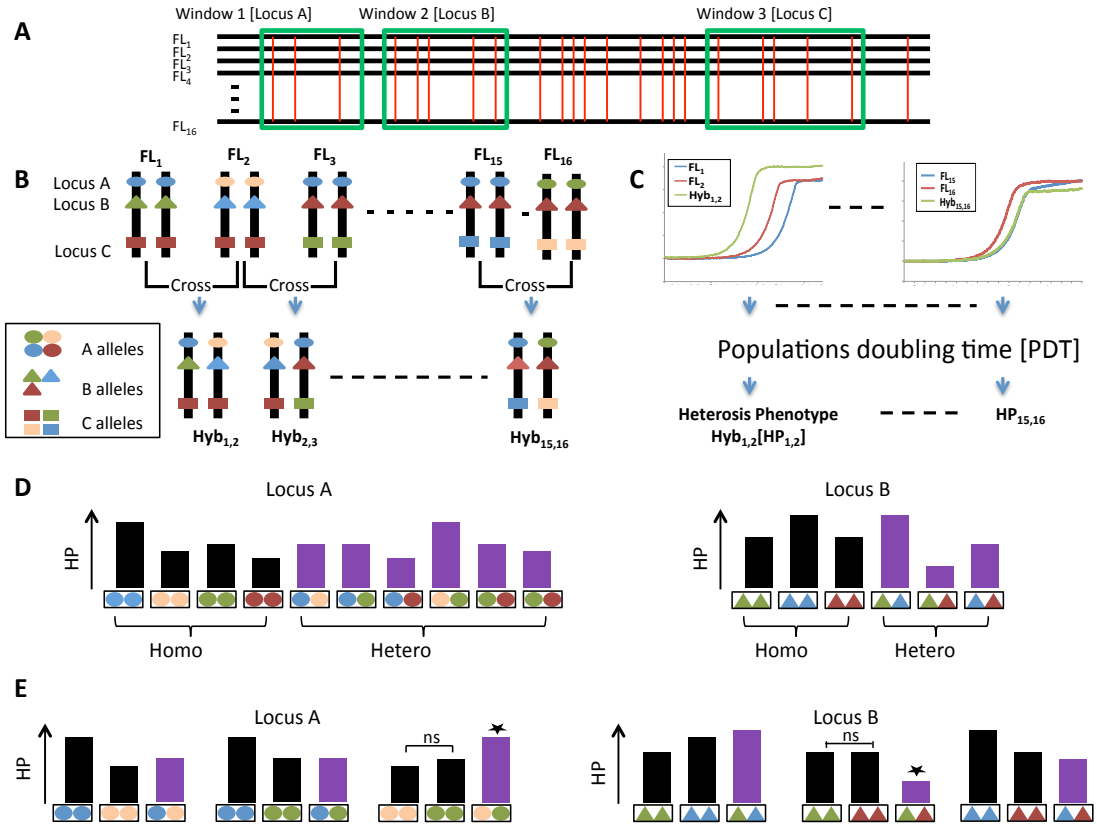

Figure 2. The HTL mapping with fully sequenced exome in *S. cerevisiae* . Genome scan of 16 *S. cerevisiae* strains following the"2x3 rule", i.e., at least two alleles, each represented by at least three parents. With minimum 3 and maximum 10 SNP per haplotype. Red lines for SNPs, green frame for windows (loci) **(B)** Allele distribution in Locus 1. Different colors represent different alleles. (the red allele is represented by only one parent, this allele will not be considered in the statistical analysis (D.2)) **(C)** Whole genome genotyping of the selected FLs followed by projecting FL genotype on hybrids and computation of their Heterotic Phenotype. Two steps of the statistical analysis: **(D)** Genomic scan for loci associated with ODH. Bars denote mean values, yet, comparison between genotypic groups was conducted between distributions (non-parameteric tests; see Methods) rather than between means or medians due to non-normal distribution of the ODH values. **(E)** Positives loci are tested for association of intra-allelic interaction with heterosis, i.e. significant difference in ODH values between hetero and two homo groups. Loci that passed these criteria (marked with *) are considered as candidate HTL. Locus 1 is underdominant HTL and locus 2 is overdominant HTL. Taken from Glikaite et al. (2015; under revision).

2.1 Heterosis dissection is sorghum populations

In this project we have used two sources for heterosis mapping, one in each country. Attempts to cross-validate results between two environments gave mixed results. This was including testing if loci associated with agronomic yield and grain nitrogen content in Israel show similar association in the sorghum association panel, or if the newly identified height QTL in the US RIL population (qHT7.1) is associated with height variation in the Israeli diallel. In both directions this comparison failed to cross-validate

results suggesting a strong genotype by environment interactions, or alternatively, different sets of alleles in the two genetic resources. Nevertheless, both groups were able to identify significant loci associated in primary populations (diallel and RIL), and validated by additional population (F2 or association panel, respectively). Both studies led to publication of the results and brief summary is given below.

2.1.1 HTL mapping in a diallel

In Israel, a panel of 273 lines was initially genotyped with 50 SSR markers and this data allowed to implement the crossing scheme and draw the principles governing such multiparental population, namely the need to balance the allelic representation to reduce to minimum the chances of identifying rare alleles across the genome. At a next step the 19 founder lines of this diallel were genotyped by sequencing at the Cornell facility and app. 20,000 SNP were included in the analysis. At this stage we have amended the HTL mapping to test for each SNP which of the modes of inheritance is more prevalent, rather than focusing solely on overdominance as originally thought in this project. This analysis led us to a very clear conclusion that overdominance is not the primary mode of inheritance governing heterosis, and instead, the model best explaining heterosis is dominant complementation. This is seen in Figure 3 that depicts the Manhattan plot for two representative traits in which only few of the loci that are associated with heterotic variation are best explained by overdominant mode of inheritance (note the number of points above the threshold of -10Log10P=6 (this was set as threshold due to multiple testing in the GWAS). Similar results were obtained for any trait considered including metabolic traits such as grain nitrogen and carbon content. More details of this part of the project are found in the annual report of 2014.

The manuscript describing these results with the SNP and grain nitrogen content data is currently being revised for submission.

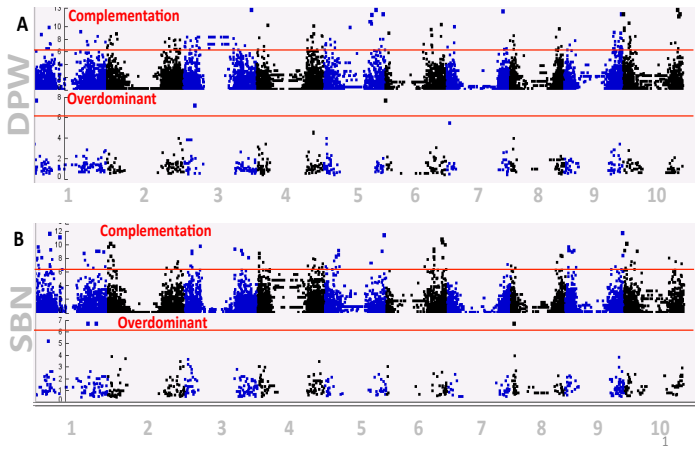

Fig. 3. Two-step genome-wide scan of HTLs for agronomic and developmental traits. **A.** Manhattan plot for any SNP significantly associated with overdominant heterosis (ODH) for dry panicle weight (DPW) regardless of mode of inheritance (top), as compared to GWAS peaks for HTLs associated with true overdominance. **B.** Similar analysis for secondary branching number of the panicle that was highly correlated with grain number(1).

2.1.1 Dissection of heterotic QTL for height

In another effort to dissect the genetic basis of heterosis in sorghum the US group has focused on plant height as the focal trait. Using a sorghum recombinant inbred lines population, a QTL (*qHT7.1*) for plant height was identified 29 cM away from the *Dw3* gene on chromosome 7. Whenever the two QTL are in repulsion linkage and two parents have opposite alleles, the hybrid could show heterosis in plant height. This was confirmed by observing plant height of hybrids crossed from RILs with different allele combinations of the two QTL (Figure 4. Alleles conferring taller plant height at each QTL is complete dominant over alleles conferring shorter plant height, agreeing with previous studies.

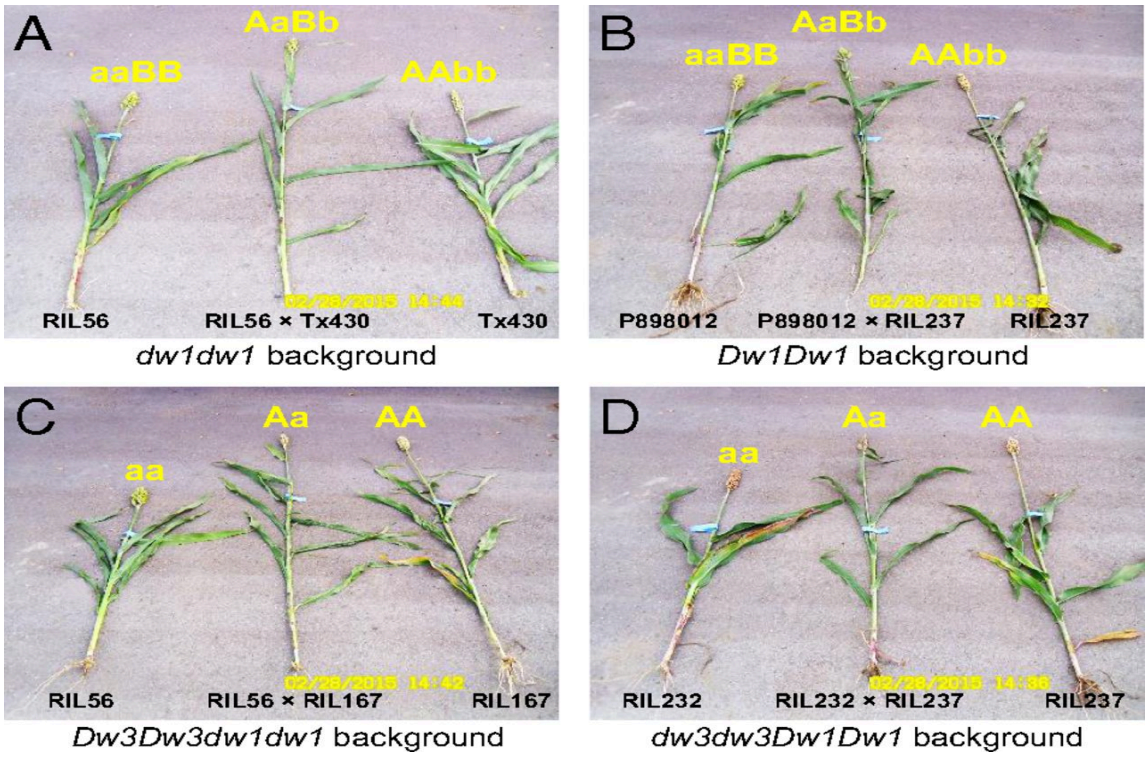

Figure 4. Plant height of parental lines and corresponding F_1 hybrids. (*A* and *B*) Heterosis caused by repulsion linkage of *qHT7.1* and *Dw3*, under different backgrounds of *Dw1*. (*C* and *D*) The tall allele of *qHT7.1* shows complete dominance over the short allele of *qHT7.1* under different backgrounds of *Dw3* and *Dw1*. In each picture, the seed parent is on the left, and the pollen parent is on the right. (Taken from Xin et al. (2015)(2))

Results from analyzing different plant height components showed that *qHT7.1* has effect on both higher and lower part of the stem, suggesting this QTL regulates plant height in a way different from *Dw3* (Figure 5*)*. Computer simulation in a segregating population showed that repulsion linkage could influence QTL detection and the estimates of additive and dominant effects. This QTL was also detected using the Sorghum Association Panel by including *Dw1*, *Dw2*, and *Dw3* genes as covariates, suggesting information from linkage mapping could guide association mapping to identify loci previously not detected(2). This dissection of the *Dw3* gene region into different QTL should enable plant breeders to fine tune plant height for grain or biomass

production, as currently being followed in the breeding program at KSU (Yu and Tesso, personal communication).

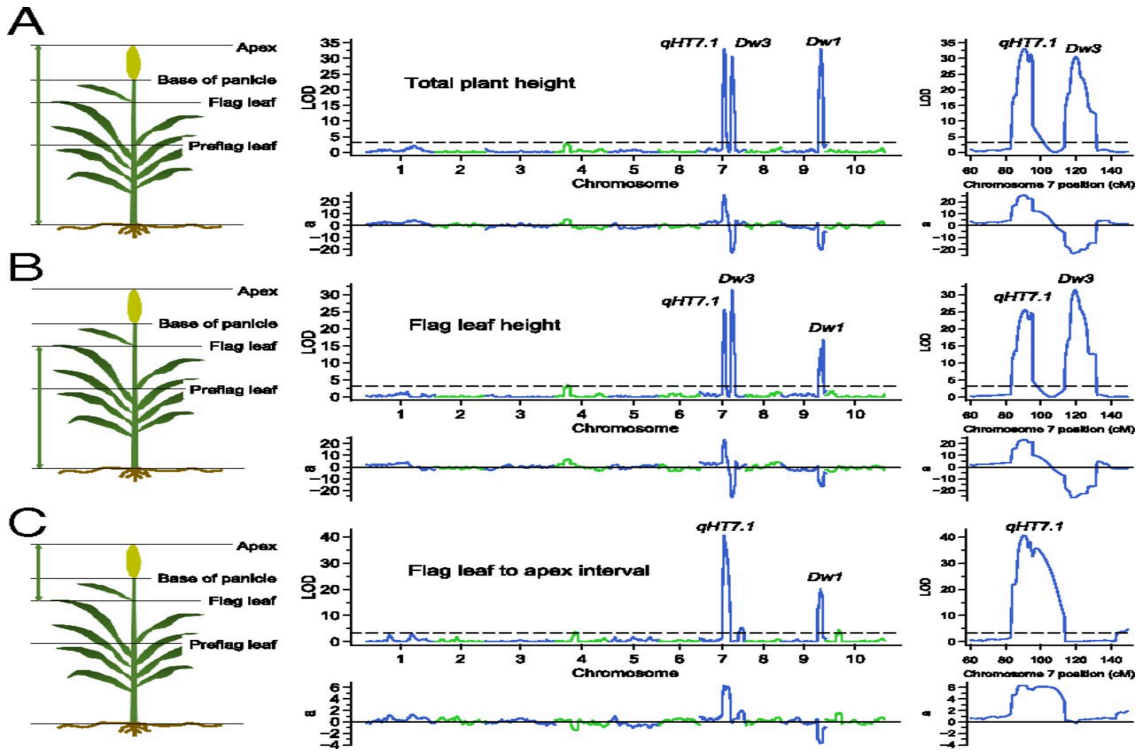

Figure 5. Linkage mapping of plant height in sorghum RIL population. (*A–C*, *Left*) Diagrams defining Results from composite interval mapping for total plant height (*A*), Flag leaf height (*B*), and Flag leaf-to-apex interval (*C*). (*Center* and *Right*) Results from composite interval mapping. (*Center*, *Upper*) The logarithm of the odds (LOD) score profile with the permutation threshold indicated by the horizontal line. (*Center*, *Lower*) The additive effect (a) with the Tx430 allele as the reference. (*Right, Upper*) The LOD score profile for enlarged chromosome 7 region. (*Right, Lower*) The additive effect

Publications for Project IS-4546-12R

Status	Type	Authors	Title	Journal	Volume: Pages	Year	Country
Published	Reviewed	*Ben-Israel,I., Kilian,B., Nida,H. and Fridman,E.*	Heterotic trait locus (HTL) mapping identifies intra-locus interactions that underlie reproductive hybrid vigor in sorghum bicolor	***PLoS One***	7 : e38993	***2012***	IS only
Published	Reviewed	*Li,X., Xianran,L., Fridman,E.., Tesso,T.T.. and Yu,J.*	Dissecting repulsion linkage in the Dw3 gene region for sorghum plant height provides insights into heterosis.	***Proc. Natl. Acad. Sci. USA***	12(38) : 11823-28	***2015***	Joint
Published	Review Article	*Fridman, E.*	Consequences of hybridization and heterozygosity on plant vigor and phenotypic stability	***Plant Science***	232 : 35-40	***2015***	IS only
Submitted	Reviewed	*Laiba, E., Glikaite ,B., Levy, Y., Paternak, Z., and Fridman, E.*	Underdominant loci associated with heterosis for population doubling time in Saccharomyces cerevisiae diallel	***Genome***	:	***2015***	IS only